

Evolutionary Trend Modeling Method of Network Event Group Based on Difference Degree

Maoyuan Zhang, Rufan Xiao, Yang Wen*, Shuyuan Sun

School of Computer Science, Central China Normal University, Wuhan Hubei
Email: *wenyang@mails.ccnu.edu.cn

Received: Oct. 6th, 2018; accepted: Oct. 16th, 2018; published: Oct. 24th, 2018

Abstract

The research of heat evolution of network events is the basis of public opinion monitoring. However, the existing research methods of network event heat mainly use the method of text processing to classify the information on the network, so as to conduct heat analysis of individual events, ignoring the correlation between events and events. This paper presents a method of modeling the evolution trend of network event groups based on the degree of difference. In this method, the heat evolution maps of multiple correlated events are pan scaled and projected into the same timeline space. The incident information is added and the matrix center is continuously updated. The objective function is minimized to find the heat evolution curve of the event group with the smallest difference degree. Experiments show that the evolution of the event group derived from this method can better reflect the uniform development characteristics of each event.

Keywords

Event Group, Heat Evolution, Degree of Difference

基于差异度的网络事件群演化趋势建模方法

张茂元, 肖茹凡, 文 洋*, 孙树园

华中师范大学计算机学院, 湖北 武汉
Email: *wenyang@mails.ccnu.edu.cn

收稿日期: 2018年10月6日; 录用日期: 2018年10月16日; 发布日期: 2018年10月24日

摘 要

网络事件的热度演化研究是舆情监测的基础, 而现有的网络事件热度研究方法主要是采用文本处理的方法。
*通讯作者。

文章引用: 张茂元, 肖茹凡, 文洋, 孙树园. 基于差异度的网络事件群演化趋势建模方法[J]. 计算机科学与应用, 2018, 8(10): 1543-1551. DOI: 10.12677/csa.2018.810168

法对网络上的信息进行归类, 从而对单个事件进行热度分析, 忽视了事件与事件间的关联。本文提出了一种基于差异度的网络事件群演化趋势建模方法。该方法将多个关联事件的热度演化图经平移缩放投影到同一时间轴空间中, 加入事件的关联信息, 不断更新矩阵中心, 最小化目标函数以求找到差异度最小的事件群的热度演化曲线。实验证明该方法得出的事件群演化曲线更能反映各个事件统一的发展特性。

关键词

事件群, 热度演化, 差异度

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在当今信息时代的大背景下, 网络事件的发生会导致民众的热烈关注, 在事件发展周期里, 民众的高度关注可以导致社会恐慌或对社会和谐构成负面影响, 因此近年党和国家对网络舆情的研究与监管也开始高度重视。目前, 国内外对于网络热点事件的研究主要集中于网络事件的应急管理和网络事件本身的客观规律两个方面, 这些研究主要的关注点都是在单个事件的发生、发展和演化机理等[1] [2] [3]。然而在如今的信息爆炸时代网络事件舆论关联是普遍存在的, 事件通过共同主体、主题或情绪等共性因素相互联系, 可以认为具有相同或相似事件原型的多个事件可构成一个事件群, 它们会因为相同或相似事件的发生而被强化和突出。

在网络事件主体关联中, 事件的舆情演化由于明显或不明显、直接或间接的主体因素互相关联, 例如陈世峰杀人事件、药家鑫事件、我爸是李刚事件的主体均为“富二代”、“官二代”, 在事件的演化周期中, 网民的舆论热点在某一时期不可避免的关注在富二代这一群体上面, 则我们认为这些网络热点事件是相互关联的。在网络事件主题关联中, 事件的舆情演化由于存在相同或相似的主题而产生联系, 比如天津港爆炸事故、日照石大化工爆炸、昆明爆炸案的主题均含有“爆炸”、“意外事故”等, 则认为这些事件是存在关联的。这些具有关联的事件我们称为事件群, 而事件群之间会产生相互影响。

基于上述思想, 本文提出了一种基于差异度的网络事件群演化趋势建模方法。该方法将多个关联事件的热度演化图经平移缩放投影到同一时间轴空间中, 加入事件的关联信息, 不断更新矩阵中心, 最小化目标函数以求找到差异度最小的事件群的热度演化曲线。实验证明该方法得出的事件群演化曲线更能反映各个事件统一的发展特性。

2. 相关研究

2.1. 事件演化相关研究

在事件演化分析研究方面, 主要的关注点都是在单个事件的发生、发展和演化机理等, 吕楠、罗军勇[2]等人提出一种基于多向量模型的事件演化分析算法, 该算法可以发现话题中各个事件间的演化关系, 解决了话题追踪技术不能对话题全局演化进行分析的问题; 张辉、李国辉[3]等人为了更好的发现话题中事件的关系, 利用事件的时间、内容、命名实体关系建立新的演化关系模型, 并提出一种新的新闻事件演化建模方法, 该方法可以准确检测事件演化并揭示事件发展的潜在关系; 廖君华、孙克迎[4]等人结合LDA主题模型和中文时间识别技术设计并实现了一个网络时序主题演化系统 HTAS, 该系统能够快速获

取和存储网络热点话题并有效分析其演化趋势; Leskovec 等人[5]提出面向话题时间序列的 K_SC 聚类算法, 它能较好地刻画话题内在发展趋势特征, 但是却无法联系事件间的关联。

2.2. 热度定义相关研究

对网络事件热度定义的研究, 主要分为两种方法: 其一是直接将事件的报道数或点击数作为热度。其二是综合考虑多方面因素定义事件的热度, 一般是将时间, 转发量, 点击率等加权求热度。例如孟琼瑶[6]在对网络热点事件热度进行预测时采用新闻的点击数和转发量作为事件的热度值; 何炎祥[7]等人将各话题的发帖数量作为话题热度, 并以此为基础改进了人口模型。郑志蕴[8]等人在内容特征、博主特征和传播特征三个方面对热门微博进行特征分解, 并使用信息增益算法对微博的热度进行度量; Pal [9]等人在 Twitter 数据集上将发帖数、回复数、转发数、粉丝数、被引用数引入热度的计算中。

3. 基于差异度的网络事件群演化趋势建模

3.1. 事件的语义指纹

当我们看见一个词语“苹果”, 并且这时看见“苹果”的解释语句“苹果是蔷薇科树的食用水果, 亦指苹果树, 果实圆形, 有红、黄或淡绿等色, 味甜或略酸”时, 我们会在脑中建立起“苹果-水果”, “苹果-圆形果实”等关系对。于是, 杨柳、何婷婷[10]等人提出了基于网络百科全书的中文词语的语义指纹概念, 用语义标签对应词或短语, 用语义指纹表示词或短语的相关词群并计算相关词群对语义标签的贡献度。通过词语的语义指纹即词语与相关词群的相关度, 人们可以很容易学习理解词或短语的意思。

根据词语的语义指纹的思想, 孟琼瑶[6]提出了事件的语义指纹的概念, 以事件为标签词, 根据事件的表示特征, 以人物、时间、地点、主要内容为语义指纹, 它们分别为事件的贡献度即语义指纹对标签词的贡献度。事件的语义指纹可以从事件的四个要素特点来描述一个事件, 与传统的文本向量相比更加完善的表达事件相关内容。

本文将事件根据事件的语义指纹表示为事件模板, 即表示事件的向量, 由事件内容特征四个方面的语义指纹对组成。表示形式为:

$$E = \{(t_1, w_1), (t_2, w_2), (t_3, w_3), \dots, (t_n, w_n)\} \quad (1)$$

其中, t 为事件的语义指纹中事件主体、主题、时间及地点的特征词, w 为特征词 t 对事件的贡献度。

3.2. 基于相似度的曲线差异度计算方法

3.2.1. 曲线的差异度计算

Leskovec 等人[11]在 2011 年提出面向话题时间序列的 K_SC 聚类算法, 它能较好地刻画话题内在发展趋势特征, 本文认为具有相同或相似事件原型的多个事件可构成一个事件群, 为了探究事件群的演化趋势, 本文加入了事件的关联因素, 提出了基于事件相似度的曲线差异度计算方法, 首先, 本文提出了两个假设, 如果两个事件热度演化图有非常相似的形状, 但在 y 轴上的投影不同, 这两个事件演化趋势仍然应该被认为是相似的。因此, 在 y 轴上缩放时间序列不应该改变相似性。其次, 即使两个事件的演化图可能会被移动, 由于他们有相似的形状其演化趋势应该被认为是相似的。因此, 转换事件演化图的时间轴不应该改变两个时间序列之间的相似性。也就是说本文认为任意两个事件的时序热度的相似性只与它们的趋势走向有关, 而和它们的峰值数值以及在何时达到峰值无关。

在事件群中, 事件和事件间具有主体、主题等共性因素, 本文将事件的相似度作为事件在事件群演化中所占权重, 将事件成功表示成事件模板向量后, 接下来就是计算事件的相似度, 本文采用基于余弦相似度的思想计算事件的相似度, 设事件 E_1 和事件 E_2 分别为:

$$E_1 = \{(t_1, w_1), (t_2, w_2), \dots, (t_n, w_n)\} \quad (2)$$

$$E_2 = \{(t_1, w_1), (t_2, w_2), \dots, (t_n, w_n)\} \quad (3)$$

则事件的相似度为:

$$Sim(E_1, E_2) = \frac{\sum_{t_i \in E_1 \cap E_2} (w_{t_i, E_1} * w_{t_i, E_2})}{\sqrt{\sum_{i=1}^{\|E_1\|} w_{t_i, E_1}^2} * \sqrt{\sum_{i=1}^{\|E_2\|} w_{t_i, E_2}^2}} \quad (4)$$

其中, $t_i \in E_1 \cap E_2$ 表示特征词 t_i 是事件 E_1 与事件 E_2 模板向量的共同元素, w_{t_i, E_1} 表示 t_i 对事件 E_1 的贡献度, w_{t_i, E_2} 表示 t_i 对事件 E_2 的贡献度, $\|E_1\|$ 和 $\|E_2\|$ 分别表示两事件 E_1 和 E_2 的模板大小。

1) 两个事件曲线的差异度计算

首先, 我们采用一个不变的距离度量公式来缩放时间序列[12]。给定两个事件 E_1 和 E_2 的热度演化时间序列 Y_1 和 Y_2 , 这两个事件的热度曲线 Y_1 和 Y_2 之间的差异度 $\hat{d}_1(Y_1, Y_2)$ 定义如下:

$$\hat{d}_1(Y_1, Y_2) = Sim(E_1, E_2) \min_{a, q} \frac{\|Y_1 - aY_{2(q)}\|}{\|Y_1\|} \quad (5)$$

其中, $Sim(E_1, E_2)$ 详见式(4), 表示两个事件 E_1 和 E_2 的相似度, $Y_{2(q)}$ 是将时间序列 Y_2 移位 q 个时间单位的结果, 在实验中, 我们将两个时间序列的第一个峰值对齐来找到 q , 而 a 为缩放的比例系数, 我们计算已知 q 的最优缩放比例系数 a 的方法为:

当 q 确定时, $\frac{\|Y_1 - aY_{2(q)}\|}{\|Y_1\|}$ 是一个关于 a 的凸函数, 因此我们可以用求偏导的方法求得最优的缩放比

例系数 a : $a = \frac{Y_1^T Y_{2(q)}}{Y_{2(q)}^T Y_{2(q)}}$, 具体推导过程如下:

令 $f = \frac{(Y_1 - aY_{2(q)})^2}{Y_1^2}$, 对 a 求偏导有:

$$\frac{\partial f}{\partial a} = -\frac{2(Y_1 - aY_{2(q)})^T Y_{2(q)}}{Y_1^2} = -\frac{2(Y_1^T Y_{2(q)} - aY_{2(q)}^T Y_{2(q)})}{Y_1^2} \quad (6)$$

令 $\frac{\partial f}{\partial a} = 0$, 可以得到:

$$Y_1^T Y_{2(q)} - aY_{2(q)}^T Y_{2(q)} = 0 \quad (7)$$

$$a = \frac{Y_1^T Y_{2(q)}}{Y_{2(q)}^T Y_{2(q)}} = \frac{Y_1^T Y_{2(q)}}{\|Y_{2(q)}\|^2} \quad (8)$$

2) 事件群曲线与多事件曲线集合之间的差异度计算

事件群可以看作是多个相关事件的集合, 我们用集合 $C = \{Y_1, Y_2, \dots, Y_n\}$ 表示多个事件的曲线集合(即事件群的曲线集合), Y_μ 表示事件群的热度演化时间序列, 则事件群曲线 Y_μ 与多事件曲线集合 C 之间的差异度 $\hat{d}_2(C, Y_\mu)$ 定义如下:

$$\hat{d}_2(C, Y_\mu) = \sum_{i=1}^n \hat{d}_1(Y_i, Y_\mu) \quad (9)$$

3.2.2. 事件群演化曲线的生成

接下来, 本文采用迭代算法的思想, 不断更新事件群的热度演化曲线, 以求找到差异度最小的事件群的演化曲线。因为本文不是求所有时间序列距离的平方和的最小值, 仅仅只更新新的事件群时间序列为所有事件演化时间序列的平均数是不合适的, 所以新的事件群曲线 Y_μ^* 应该是在所有热度演化时间序列集合 C 上的 $\hat{d}_2(C, Y_\mu)^2$ 的和的最小值, 即定义新的事件群的热度演化时间序列 Y_μ^* 如下:

$$Y_\mu^* = \arg \min \sum_{i=1}^n \hat{d}_2(Y_i, Y_\mu)^2 \quad (10)$$

由于在平移步骤中找到了最佳平移系数 q_i , 然后用最优缩放比例系数代入 a_i , 得到:

$$Y_\mu^* = \arg \min \frac{1}{\|Y_\mu\|^2} \sum_{i=1}^n \left\| \frac{Y_i^T Y_\mu}{\|Y_i\|^2} Y_i - Y_\mu \right\|^2 \quad (11)$$

化简后可以得到:

$$Y_\mu^* = \arg \min \frac{1}{\|Y_\mu\|^2} \sum_{i=1}^n \left\| \frac{Y_i Y_i^T Y_\mu}{\|Y_i\|^2} - Y_\mu \right\|^2 = \arg \min \frac{1}{\|Y_\mu\|^2} \sum_{i=1}^n \left\| \left(\frac{Y_i Y_i^T}{\|Y_i\|^2} - I \right) Y_\mu \right\|^2 = \arg \min \frac{1}{\|Y_\mu\|^2} Y_\mu^T \sum_{i=1}^n \left(I - \frac{Y_i Y_i^T}{\|Y_i\|^2} \right) Y_\mu \quad (12)$$

最后, 用 M 代替 $\sum_{i=1}^n \left(I - \frac{Y_i Y_i^T}{\|Y_i\|^2} \right)$ 得到 Y_μ^* , 如下:

$$Y_\mu^* = \arg \min \frac{Y_\mu^T M Y_\mu}{\|Y_\mu\|^2} \quad (13)$$

如果我们通过乘以 M 的特征向量转换 Y_μ , $Y_\mu^T M Y_\mu$ 相当于 M 的特征值的加权和, 其最小值是 $\lambda_m \|Y_\mu\|^2$, λ_m 是特征向量 Y_{μ_m} 到矩阵 M 的最小特征值 λ_m 。因此, 公式(12)是 λ_m 让 $Y_\mu = Y_{\mu_m}$ 达到最小值。由于 M 是由时间序列 Y_i 得出的, 我们可以说为新的矩阵中心(即事件群曲线) Y_μ^* 找到最小的特征向量 M 。

4. 实验与分析

4.1. 实验数据

本章实验数据来自爬取新浪新闻中心的历年新闻语料经处理后得到的新闻事件, 本实验选取了 18 个事件群, 共计 2012 篇新闻文本作为数据集, 事件群及其包含事件列表如表 1。

4.2. 评价指标与对比方法

本文的研究目标是得到事件群的演化趋势图, 最终的结果与单个事件的演化图差异应该越小越好, 因此我们用 F -Value(F 值)来作为本章实验的评价指标, F 值反映了事件群中与各个事件热度演化的紧凑程度, F 值越小表明越紧凑差, 异越小。其计算方法如下(公式):

$$F = \sum_{i=1}^n \min_{a, q} \frac{\|Y_i - a Y_{\mu(q)}\|^2}{\|Y_i\|^2} \quad (14)$$

其中, Y_i 是事件群中某个事件的热度演化序列, Y_μ 是得到的事件群的热度演化序列。

为了验证本章实验的有效性, 将本章提出的方法与其他两种方法进行比较。

方法 1: 算术平均, 事件群中几个事件的热度算术平均得到事件群演化结果。

Table 1. Network event groups and event lists
表 1. 网络事件群及事件列表

| 事件群 | 事件 |
|--------|--|
| 暴力恐怖 | 3.1 昆明火车站暴力恐怖袭击事件、4.30 乌鲁木齐火车站爆炸事件、12.28 墨玉县暴恐袭击事件、7.18 新疆和田暴恐事件、10.28 金水桥暴力恐怖袭击事件 |
| 地震 | 5.12 汶川地震、4.14 玉树地震、4.20 雅安地震、1.12 南黄海地震 |
| 恶意杀人 | 4.1 复旦投毒案、5.12 陕西幼儿园惨案、6.22 杭州小区纵火案、11.3 留日女生遇害案 |
| 火灾 | 香格里拉大火事件、青岛化工厂爆炸事故、丽江古城火灾事故、1.2 兰州火灾事故 |
| 艺人吸毒 | 8.17 柯震东吸毒被抓事件、3.17 李代沫吸毒被抓、酒井法子吸毒事件、房祖名涉毒案 |
| 虐待幼童 | 携程幼儿园虐童事件、北京红黄蓝虐童事件、6.21 南京女童饿死案、郑州电梯保姆虐童事件 |
| 暴雨洪水灾害 | 7.19 河北特大暴雨洪涝灾害、9.13 上海特大暴雨、7.23 武汉特大暴雨 7.21 北京特大暴雨 |
| 食品安全 | 日本毒大米事件、三聚氰胺毒奶粉事件、蒙牛致癌门、9.4 台湾地沟油事件 |
| 名人酒驾 | 高晓松醉驾案、郎永淳酒驾案、王志文酒驾被抓 |
| 交通事故 | 6.20 南京重大交通事故、温州动车事故、3.16 山西临汾特大交通事故、7.19 沪昆高速客货车相撞事故 |
| 教育管理 | 安徽高考眼镜门事件、湖北江苏高考减招风波、中关村二小疑似校园欺凌事件、河南女孩王娜娜被顶替上大学 |
| 医患矛盾 | 东北女孩怒斥广安门医院号贩子、安徽男子右肾失踪风波、北医三院产妇事件、广东省人民医院退休主任被捅事件 |
| 医药安全 | 山东问题疫苗案、毒胶囊事件、甘肃陇西硫磺熏药材事件、 <u>维 C 银翘片含毒案件</u> |
| 国家安全 | 中日钓鱼岛争端、香港占中事件、美国拟在韩国部署萨德系统、南海仲裁案 |
| 涉法涉诉 | 9.10 代理书记黄兴国被查、8.26 甘肃白银连环杀人案告破、6.3 广西律师遭法警殴打、11.20 内蒙呼格案、5.7 雷洋案件 |
| 旅游管理 | 哈尔滨天价鱼事件、10.4 青岛大虾事件、三亚宰客门、2.4 丽江女游客被打事件 |
| 公共安全 | 4.3 北京和颐酒店劫持事件、8.9 重庆女大学生搭错车遇害案、8.21 女大学生济南失联、8.28 江苏 19 岁女大学生失踪事件 |
| 网络治理 | 魏则西之死、山东徐玉玉电信诈骗案、大学生校园贷事件、11.30 罗一笑事件 |

方法 2：相似度加权平均，将事件群中几个事件的相似度作为权重对热度进行加权平均得到事件群演化结果。

4.3. 实验结果分析

以暴力恐怖事件群为例，其由 5 个主题相关事件组成，分别是 3·1 昆明火车站暴力恐怖袭击事件、4·30 乌鲁木齐火车站爆炸事件、12·28 墨玉县暴恐袭击事件、7·18 新疆和田暴恐事件、10·28 金水桥暴力恐怖袭击事件，以新闻报道的评论数作为热度，按天进行划分，选取从事件发生开始的前十五天的热度值。关联事件的事件热度演化趋势如图 1 所示。

在本实验中得到的事件群热度演化曲线结果如图 2 所示，可以看出这些关联事件构成的事件群的演化趋势是一个快速上升到达波峰后开始下降并伴随波动逐渐趋于平稳的走势，并且事件群的演化趋势对应于单个事件的走势基本相同(对应图 1 的(a)~(e))。对于暴力恐怖事件群的热度走势分析，是由于在事件刚发生时报道量和网民关注度激增，此时报道主要关注在恐怖袭击造成的人员伤亡和恐怖分子的残暴行为从而带动网民的愤怒、惊恐等情绪，随着报道量和关注度的增加事件热度会在 3 至 4 天左右达到峰值，之后随着恐怖分子被抓获及政府有关部门的积极处理，网民的关注度会逐渐减弱，事件的热度也会逐渐趋于平稳。

表 2 给出了 18 个事件群演化曲线在三种方法中的 F 值(越小越好)，从表 2 可以看出本文提出的基于

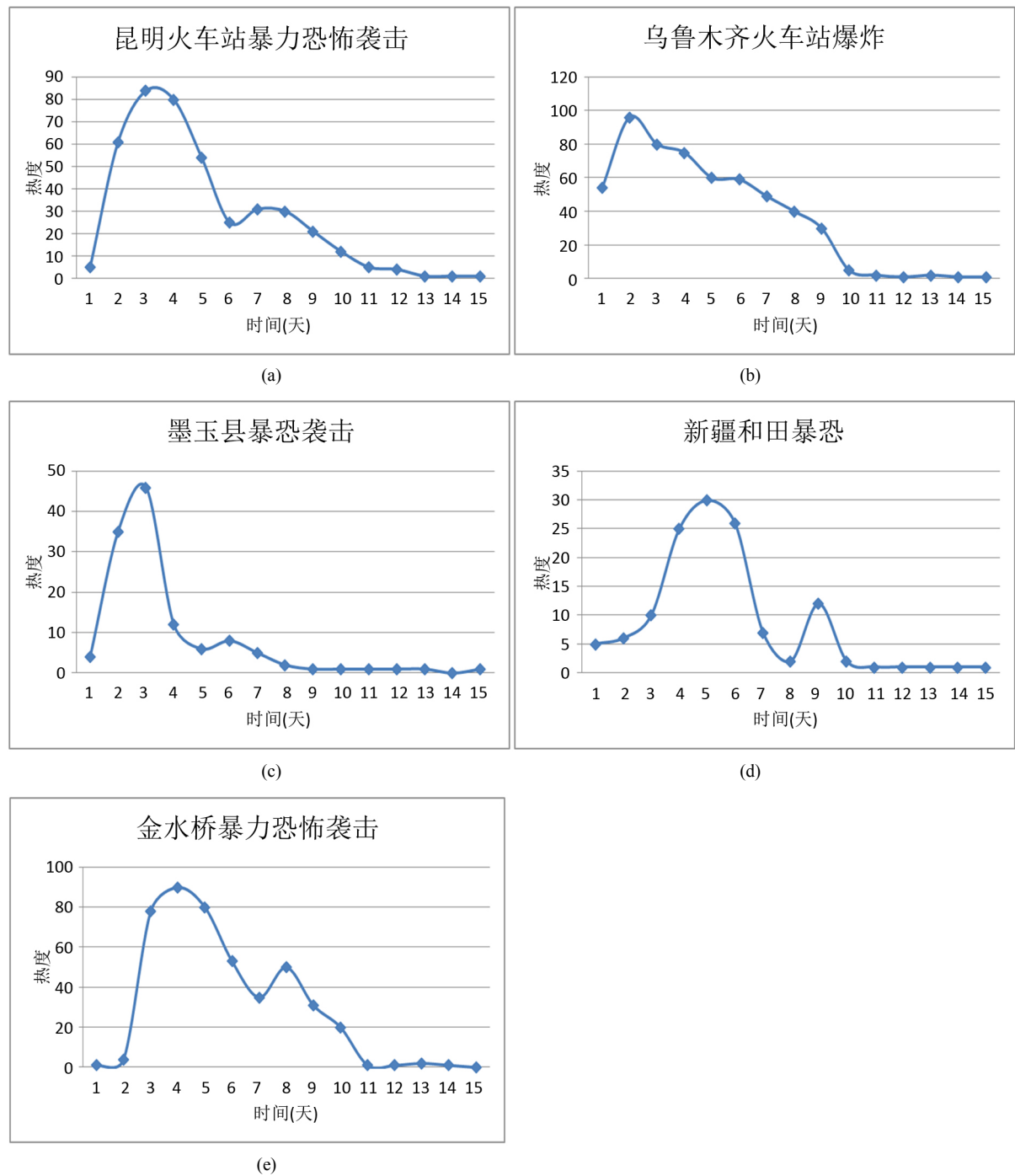


Figure 1. Heat evolution map of event

图 1. 事件的热度演化图

相似度的曲线差异度计算方法在 F 值上优于另外两个求事件群热度曲线的方法, 所以本文提出的方法得出的事件群演化曲线更能反映各个事件统一的发展特性。

5. 结束语

本文基于事件与事件间存在关联的思想提出了一个网络事件群演化趋势建模方法。该方法将多个关

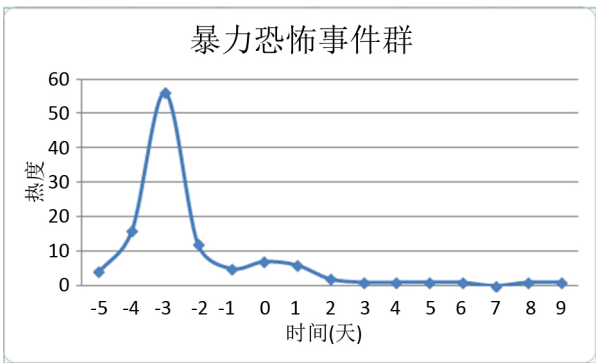


Figure 2. The heat evolution map of the violent terrorist event group

图 2. 暴力恐怖事件群的热度演化图

Table 2. F values of three methods under different event groups

表 2. 三个方法在不同事件群下的 F 值

| 事件群 方法 | 基于相似度的曲线差异度计算方法 | 算术平均曲线计算方法 | 相似度加权平均曲线计算方法 |
|-----------|-----------------|------------|---------------|
| 暴力恐怖 | 39.68 | 59.41 | 52.36 |
| 地震 | 42.41 | 61.72 | 53.16 |
| 恶意杀人 | 67.50 | 85.42 | 79.37 |
| 火灾 | 37.24 | 50.11 | 43.54 |
| 艺人吸毒 | 48.36 | 60.45 | 61.51 |
| 虐待幼童 | 56.61 | 64.37 | 71.92 |
| 暴雨洪水灾害 | 49.91 | 69.88 | 65.40 |
| 食品安全 | 52.47 | 61.47 | 59.61 |
| 名人酒驾 | 36.69 | 46.17 | 47.95 |
| 交通事故 | 61.44 | 79.55 | 71.10 |
| 教育管理 | 39.95 | 55.98 | 51.75 |
| 医患矛盾 | 55.53 | 71.43 | 67.26 |
| 医药安全 | 34.32 | 52.62 | 46.80 |
| 国家安全 | 41.51 | 59.71 | 55.29 |
| 涉法涉诉 | 72.42 | 79.91 | 81.13 |
| 旅游管理 | 68.13 | 75.44 | 73.31 |
| 公共安全 | 47.12 | 61.56 | 58.24 |
| 网络治理 | 69.53 | 78.15 | 74.37 |

联事件的热度演化图经平移缩放投影到同一时间轴空间中，加入事件的关联信息，不断更新矩阵中心，最小化目标函数以求找到差异度最小的事件群的热度演化曲线。实验证明该方法得出的事件群演化曲线更能反映各个事件统一的发展特性。找到事件群的演化趋势结果，当发生类似事件时我们可以通过事件群的演化趋势判断事件的演化趋势，可以更好地帮助政府有关部门把控网络热点事件的舆情发展态势，提高网络舆情监管功效。

基金项目

教育部人文社会科学研究基金项目(项目批准号: 15YJC870029); 国家语委科研项目(No.YB135-40); 华中师范大学中央高校基本科研业务费项目(No.CCNU16GD010)。

参考文献

- [1] Nan, L.V., Luo, J.Y., Liu, Y., *et al.* (2009) Effective Event Evolution Analysis Algorithm. *Application Research of Computers*, **26**, 4101-4103.
- [2] 吕楠, 罗军勇, 刘尧, 等. 一种有效的事件演化分析算法[J]. 计算机应用研究, 2009, 26(11): 4101-4103.
- [3] 张辉, 李国辉, 孙博良, 等. 一种新闻事件演化建模方法[J]. 国防科技大学学报, 2013, 35(4): 166-170.
- [4] 廖君华, 孙克迎, 钟丽霞. 一种基于时序主题模型的网络热点话题演化分析系统[J]. 图书情报工作, 2013, 57(9): 96-102, 118.
- [5] Leskovec, J., Backstrom, L. and Kleinberg, J. (2009) Meme-Tracking and the Dynamics of the News Cycle. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, 497-506.
- [6] 孟琼瑶. 基于语义指纹关联的网络事件热度分析方法研究[D]. 华中师范大学, 2016.
- [7] 何炎祥, 刘健博, 刘楠. 基于改进人口模型的微博话题趋势预测[J]. 通信学报, 2015(4): 5-12.
- [8] 郑志蕴, 江国林, 张行进, 等. 基于多特征的热门微博预测算法研究[J]. 小型微型计算机系统, 2017(3): 494-498.
- [9] Pal, A. and Counts, S. (2011) Identifying Topical Authorities in Microblogs. *Proceedings of the 4th ACM international Conference on Web Search and Data Mining*, New York, 45-54.
- [10] 杨柳, 何婷婷, 涂新辉. 基于网络百科全书的中文关联语义知识获取[C]//第五届全国青年计算语言学研讨会论文集. 2010.
- [11] Yang, J. and Leskovec, J. (2011) Patterns of Temporal Variation in Online Media. *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, ACM, 177-186.
- [12] Chu, K.K.W. and Wong, M.H. (1999) Fast Time-Series Searching with Scaling and Shifting. *Proceedings of the Eighteenth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, ACM, 237-248.
<https://doi.org/10.1145/303976.304000>

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2161-8801, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: csa@hanspub.org